PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04N 5/14 | A1

(11) International Publication Number: **WO 98/23085**

(43) International Publication Date: 28 May 1998 (28.05.98)

(21) International Application Number: PCT/CA97/00881

(22) International Filing Date: 20 November 1997 (20.11.97)

(30) Priority Data:
2,190,785 20 November 1996 (20.11.96) CA

(71) Applicant *(for all designated States except US)*: TELE-VITESSE SYSTEMS INC. [CA/CA]; 115 rue Champlin, Hull, Quebec J8X 3R1 (CA).

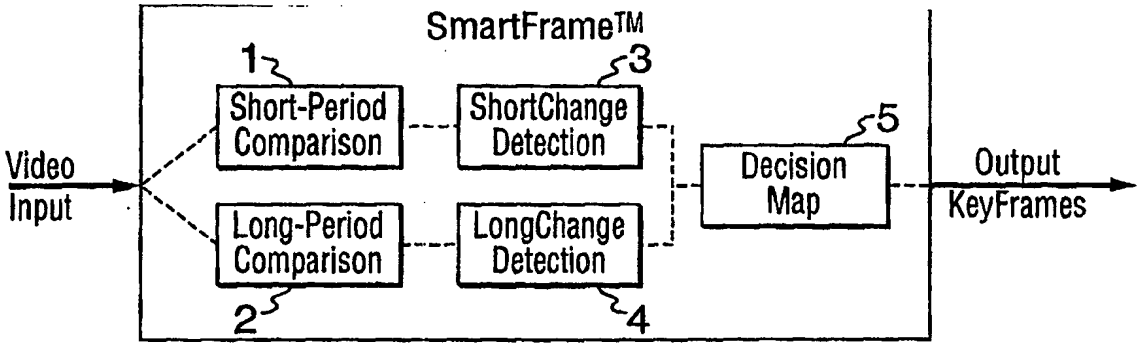
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: HIRZALLA, Nael [CA/US]; Apartment 1077, 3503 S. Lamar, Austin, TX 78704 (US). STREATCH, Paul [CA/CA]; P.O. Box 1196, Richmond, Ontario K0A 2Z0 (CA). MACLEAN, Roger [CA/CA]; 90 St. Claire Avenue, Nepean, Ontario K2G 2A6 (CA). MENARD, Rob [CA/CA]; 168 Oakridge Boulevard, Nepean, Ontario K2G 2V2 (CA).

(74) Agent: MITCHELL, Richard, J.; Marks & Clerk, P.O. Box 957, Station B, Ottawa, Ontario K1P 5S7 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: METHOD OF PROCESSING A VIDEO STREAM



(57) Abstract

A method of processing a video stream, involves contemporaneously selecting first and second pairs of frames in the video stream with a predetermined period. The second pairs of frames have a longer period than the first pairs. For each of the first and second pairs of frames, there is determined a difference value representing the number of pixels whose value has changed between the first and second frames of the pair. A particular logic level is generated depending on whether this difference value exceeds a predetermined threshold. The generated logic levels are then compared with a decision map to identify cuts in the video stream.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

METHOD OF PROCESSING A VIDEO STREAM

This invention relates to a method of processing a video stream, to detect changes, for example, a cut in scenes.

In video terminology, a video stream consists of a number of frames that are displayed successively to create the illusion of motion. A sequence of frames can be considered to form a "scene", which is considered to be a continuous action in space and time (i.e. with no camera breaks). A "cut" is a discontinuity between scenes. A cut is sharp if it can be located between two frames and gradual if it takes place over a sequence of frames. A keyframe is a frame that represents a whole scene. It can either be calculated or selected from the frames of the scene it represents.

There are many situations where it is desirable to select a cut. For example, selecting keyframes to transmit over a network, save onto a hard disk, or use to browse a video can reduce bandwidth, capacity and time than considering the whole video data. However, video segmentation is a difficult process in view of the various types of camera breaks and different operations that can take place.

Video parameters include intensity, red-green-blue (RGB), hue-value-chroma (HVC), and a motion vector. A traditional approach for detecting a cut is to compare one or more of these parameters, such as intensity, of the corresponding pixels in a pair of consecutive frames. If the number of pixels whose intensity values have changed from one frame to the next exceeds a certain threshold, a cut is presumed. However, such an approach results in low detection rates and result in the detection of false cuts or missing real cuts. False cuts may result from camera operations, object movements or flashes within a video clip, while missed cuts may result from gradual scene changes.

An object of the invention is to alleviate the afore-mentioned problems.

According to the present invention there is provided a method of processing a video stream, comprising the steps of selecting first pairs of frames in the video stream with a predetermined period; contemporaneously selecting second pairs of frames in the video stream, said second pairs of frames having a longer period than said first pairs of frames; for each of said first and second pairs of frames, determining a difference value

representing the degree of change between the first and second frames of the pair and generating a particular logic level depending on whether this difference value exceeds a predetermined threshold; and comparing the generated logic levels with a decision map to identify cuts in the video stream.

The degree of change may be represented by the number of pixels for which a particular value, such as intensity, has changed. Alternatively, the difference value may be arrived at by, for example, taking the root mean square of the differences in pixel values. In this case, the difference in intensity value of each corresponding pair of pixels is determined, the results squared, and the square root taken of the sum. This rms value can then be compared to a threshold. A value other than intensity, for example hue, can be chosen for the value.

By this method, gradual cuts between scenes can be more accurately detected and the occurrence of false detections can be reduced.

In a preferred embodiment, , the change in difference value between each of the first and second pairs of frames and the corresponding previous pairs is determined, and additional logic levels are generated that depend on whether the change in difference values exceeds a predetermined threshold. The additional logic levels are also compared with the decision map to assist in identifying the cuts. This additional step enhances the detection process.

The invention also provides video processing apparatus comprising means for selecting first pairs of frames in the video stream with a predetermined period; means for contemporaneously selecting second pairs of frames in the video stream, said second pairs of frames having a longer period than said first pairs of frames; means for determining, for each of said first and second pairs of frames, a difference value representing the degree of change between the first and second frames of the pair and generating a particular logic level depending on whether this difference value exceeds a predetermined threshold; and means for comparing the generated logic levels with a decision map to identify cuts in the video stream.

The invention will now be described in more detail, by way of example, only with reference to the accompanying drawings, in which:-

Figure 1 is a block diagram of an apparatus for detecting cuts in a video stream;

Figure 2 illustrates the main processing routine;

Figure 3 illustrates the detection processing routine;

Figure 4 illustrates the compare frames processing routine;

Figure 5 illustrates the change detection routine;

Figure 6 illustrates the set tag routine; and

Figure 7 shows typical threshold values.

The system illustrated is typically implemented on a Pentium 133 MHz computer. A digital video stream, for example, from a digital video camera or an analog feed passed through a digital-to-analog converter (not shown), is split and passed to short and long period units 1, 2. The short period comparison unit identifies a pair of frames in a stream, for example, the fourth and fifth frames, and determines the number of pixels whose values have changed. This number is then compared with a threshold and allocated a logic level 1 if it exceeds the threshold and otherwise a logic level 0. The pixel values can be any suitable value, but typically the intensity is used.

The long period comparison unit 2 carries out the same operation, except on pairs of frames that are further apart, for example, first and eighth frames in the video stream. It generates a logical 1 if the number of pixels whose intensity values have changed exceeds a predetermined threshold. Otherwise it generates a logical 0.

The video stream is then passed to the short period change detection unit 3 and the long period change detection unit 4. The short period change detection unit 3 compares the current interframe difference value, derived in unit 1, namely the number of pixels whose intensity values have changed between each pair of pixels, with the previous pair, or the average of all the previous pairs, of interframe difference values to derive the change. If the change in interframe difference values exceeds a predetermined threshold, a logical 1 is generated, otherwise a logical 0 is generated.

The long period change detection unit 4 does the same as the short period change detection unit, except with frames separated by a longer period, the first and eighth frames

in this example. The threshold for the long period change detection unit is typically higher than for the short period detection unit.

A decision map shown below is stored in unit 5.

Short-Change	Short-Period Cut	Long-Change	Long-Period Cut	Type of Change	Cut?
0	X	X	0	-	No
X	X	0	1	*Action*	No
X	0	1	1	**Gradual**	Yes
0	1	1	1	*Action*	No
1	X	X	0	Flashes	No
1	1	1	1	**Sharp**	Yes

This contains a table of all possible values of the logic outputs of units 1 to 4, logic level 0 representing a comparison below the threshold, logic 1 being a comparison above the threshold and X, where X means "don't care". i.e. 0 or 1. For example, in the short change column, a 1 means that unit 3 detection a change in difference values between successive pairs of frames above a threshold, 0 means any change was below the threshold, and X means that the outcome of the short change comparison is not relevant to the decision.

The system shown in Figure 1 moves through successive frames as follows. For example, if the system processes six frames at a time, frames 1 and 6 would form the long pair and frames 3 and 4 might form the short pair. If no cuts are detected in this block, the next block will be made of frames 2 to 7, with frames 2 and 7 forming the long pair, and frames 4 and 5 forming the short pair and so on. However, if a cut is detected, the next block will contain frames 6 to 11 since the block size is very small compared to a typical scene length and no two cuts can be detected within one block.

On looking at the table above, it will be observed that a positive result for all four comparisons indicates a true sharp cut, whereas a positive result in the long change and long period detector without a corresponding result in the short period detector indicates a probable gradual cut.

A frame may contain one or more windows to which the above described process can be applied. The use of windows allows the system to focus on specific areas of a frame instead of considering the frame as a whole. This can save considerable processing time.

The main processing routine is shown in Figure 2. Starting at block 10, the routine determines at process block11 whether the frame is decompressed. If not, it is decompressed in process block 12. At process block 13, the frame number is incremented. Decision block 14 determines whether variable *jump* (set in the Tag routine to be described) is greater than 0 and the number of sectors is greater than 1. If not, and the *FrameNo* variable is greater than the *Blocksize* variable, block 15 calls the call detection routine shown in Figure 3. If the output of Decision block 14 is false, block 16 decrements the *jump* variable.

Figure 3 shows the cut detection routine. Starting from process block 21, block 22 performs the short frame comparison and block 23 performs the long frame comparison to determine the number of pixels whose intensity values have changed for each pair frames.

Process blocks 27, 28 determine whether the percentage change is above the threshold for the associated window, and if so generate a logical 1, otherwise they generate a logical zero.

Process blocks 25 and 26 perform the short and long change detection operations following which block 29 calls the set tag routine described with reference to Figure 4. Process block 30 causes the routine to loop for all windows of interest assuming the system is operating on the basis that there is more than one window. Of course, there could only be one window that represents a whole frame.

Figure 4 shows the compare frames routine in more detail. Processing blocks 40, 41 loop for all the pixels in a row and column respectively. Block 42 then finds the differences between the corresponding pixels in the first and second frames of each pair. If the decision unit 43 finds a difference greater than a threshold, the *CP* variable, which represents the percentage change of the window containing the pixels is incremented.

Figure 5 shows the change detection routine identified in blocks 25, 26 in Figure 3 in more detail. Block 50 finds the difference between the change period *CP* for the window and the average change period *CPA*. If this change is greater than a threshold, as determined in decision unit 53, and the condition in block 54 is met, process block 55 sets the output to logical 1 to indicate a change. The change detection routine shown in Figure 5 works for both long period and short period changes.

Referring now to Figure 6, the tag routine essentially implements the decision map table shown in Figure 7. Starting from process block 60 called by process block 29 in Figure 6, the routine determines whether there has been a short period change in block 61. If yes, decision block 63 determines whether there has been a short period cut, long period cut, and long period change. If yes, the block 67 creates a cut tag. If no, the block 68 determines whether there has been long period cut. If yes, block 71 creates a flasg tag.

If the result of decision block 68 is negative, block 69 checks the condition *!SPcut*, and *LPcut* and *LPchange* is met, where !SPcut means that the SPcut variable is logical 0, or in other words no cut was detected in the short period. If yes, block 73 creates a cut tag. If no, block 74 creates an action tag.

If the result of decision block 61 is no, decision block 62 determines whether the condition *!SPcut* and *LPcut* and *Lpchange* has been met. If yes, block 64 creates a cut tag and sets the variable *jump* equal to the block size. If no, block 65 determines whether there has been an LP cut. If no, block 70 creates a no cut tag; if yes, block 74 creates an action tag.

The program then moves on to a following block of frames to repeat the process, continually creating tags identifying cuts and indicating whether a cut has been detected in the processed block.

Figure 7 shows typical short-period interframe difference values expressed as a percentage vs. Frame number. Although $Threshsold_{LONGPERIOD}$ is applied on the corresponding long-term chart, it is shown in this figure as well.

The described method has many uses. For example, it can be applied to scene change detection with automatic cut detection and flagging, visual index creation for

videos, video transformation from video to pictorial transcripts and illustrated audio, video sequence recognition, motion detection, motion tracking and sizing, and bandwidth reduction by extracting only changing information from a scene.

The described method can achieve a high and robust video cut detection rate in part due to the change detection routine, satisfy real-time requirements. It can easily be applied only to specific windows of interest within a frame in the manner described. It can be applied to automatic television monitoring and be situated either at the network access point or at the user end. It can also be integrated with any database management system that needs to index or store video.

Claims:

1. A method of processing a video stream, comprising the steps of:

selecting first pairs of frames in the video stream with a predetermined period;

contemporaneously selecting second pairs of frames in the video stream, said second pairs of frames having a longer period than said first pairs of frames;

for each of said first and second pairs of frames, determining a difference value representing the degree of change between the first and second frames of the pair and generating a particular logic level depending on whether this difference value exceeds a predetermined threshold; and

comparing the generated logic levels with a decision map to identify cuts in the video stream.

2. A method as claimed in claim 1, wherein the degree of change is represented by the number of pixels for which a value has changed.

3. A method as claimed in claim 1 or claim 2, wherein said threshold is different for each of said first and second pairs of frames.

4. A method as claimed in any one of claims 1 to 3, further comprising, for each of said first and second pairs of frames, determining the change in the interframe difference value for successive pairs of frames, and generating additional logic levels depending on whether the change in difference values exceeds a predetermined threshold, said additional logic levels also being compared with said decision map to assist in identifying said cuts.

5. A method as claimed in claim 4, wherein said change in interframe difference value is determined by comparing the interframe difference for the current pair of frames with the average interframe difference for previous pairs of frames.

6. A method as claimed in any of claims 1 to 5, wherein said value is the pixel intensity.

7. A method as claimed in any of claims 1 to 6, wherein said frames are divided into at least one window and said processing steps are carried out within a selected window.

8. A method of detecting scene changes in a video stream comprising detecting cuts by a method as claimed in any one of claims 1 to 7.

9. A method of creating a video index which includes detecting cuts by a method as claimed in any one of claims 1 to 7.

10. A method of transforming from video to pictorial transcripts which includes detecting cuts by a method as claimed in any one of claims 1 to 7.

11. A method of recognizing video sequences which includes detecting cuts by a method as claimed in any one of claims 1 to 7.

12. A method of motion detection which includes detecting cuts by a method as claimed in any one of claims 1 to 7.

13. A method of motion tracking which includes detecting cuts by a method as claimed in any one of claims 1 to 7.

14. A method of bandwidth reduction which includes detecting cuts by a method as claimed in any one of claims 1 to 7.

15. Video processing apparatus comprising:

means for selecting first pairs of frames in the video stream with a predetermined period;

means for contemporaneously selecting second pairs of frames in the video stream, said second pairs of frames having a longer period than said first pairs of frames;

means for determining, for each of said first and second pairs of frames, a difference value representing the degree of change between the first and second frames of the pair and generating a particular logic level depending on whether this difference value exceeds a predetermined threshold; and

means for comparing the generated logic levels with a decision map to identify cuts in the video stream.

16. Video processing apparatus as claimed in claim 15, wherein said determining means determines the number of pixels for which a value has changed.

17. An apparatus as claimed in claim 14 or 15, further comprising, for each of said first and second pairs of frames, means for determining the change in the interframe difference value for successive pairs of frames, means for generating additional logic levels depending on whether the change in difference values exceeds a predetermined threshold, and said comparing means also comparing said additional logic levels with said decision map to assist in identifying said cuts.

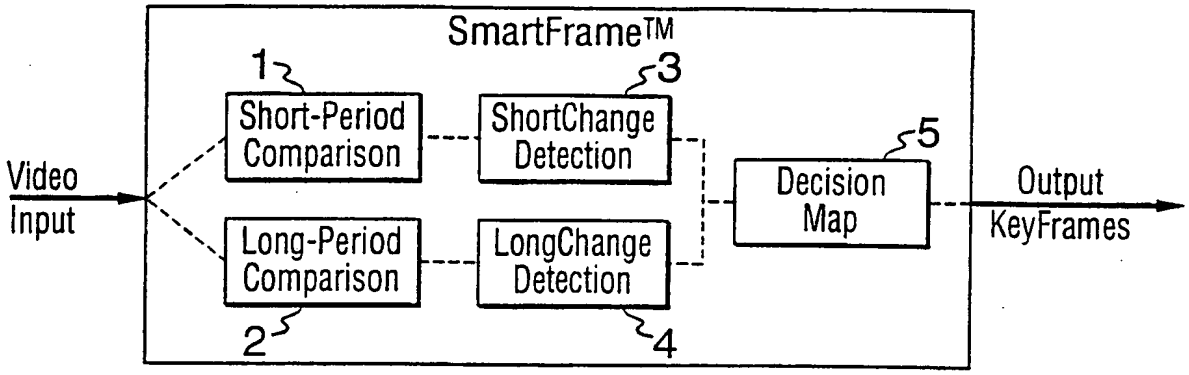

SmartFrame™
1
Short-Period Comparison
3
ShortChange Detection
5
Decision Map
Video Input
Output KeyFrames
Long-Period Comparison
LongChange Detection
2
4

FIG. 1

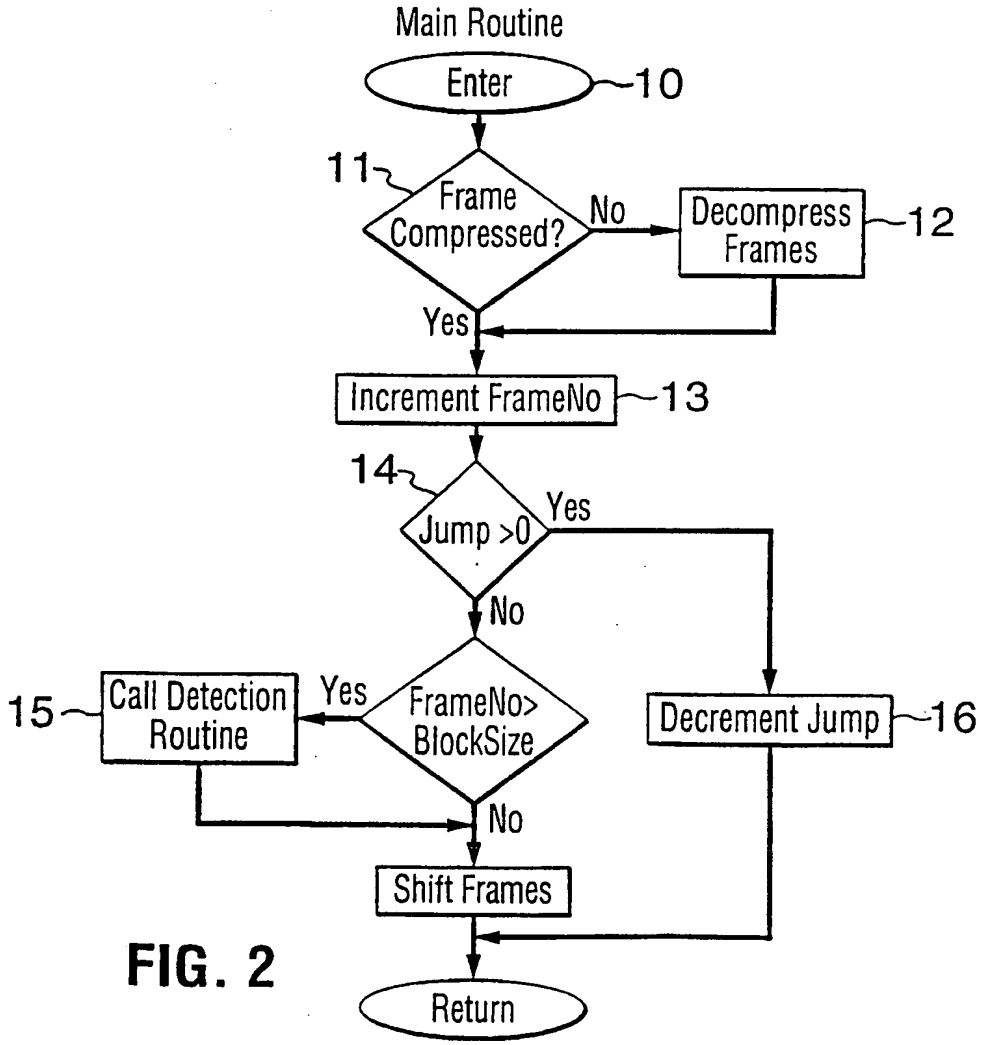

Main Routine
Enter
10
11
Frame Compressed?
No
Decompress Frames
12
Yes
Increment FrameNo
13
14
Jump >0
Yes
No
15
Call Detection Routine
Yes
FrameNo> BlockSize
Decrement Jump
16
No
Shift Frames
Return

FIG. 2

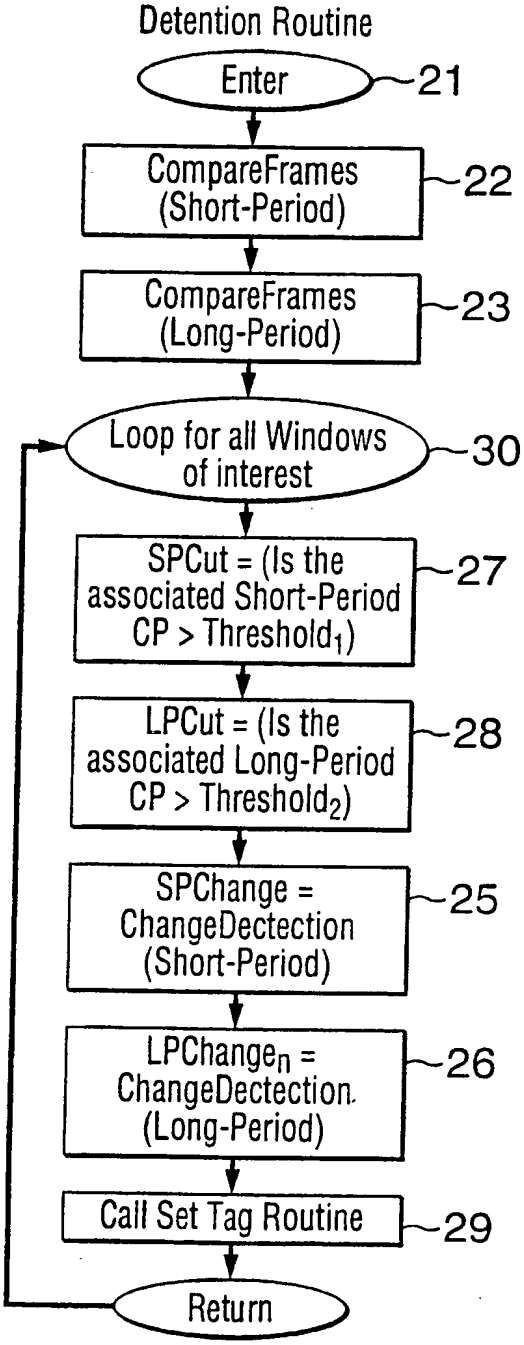

Detention Routine
Enter
21
CompareFrames (Short-Period)
22
CompareFrames (Long-Period)
23
Loop for all Windows of interest
30
SPCut = (Is the associated Short-Period CP > Threshold1)
27
LPCut = (Is the associated Long-Period CP > Threshold2)
28
SPChange = ChangeDectection (Short-Period)
25
LPChangen = ChangeDectection (Long-Period)
26
Call Set Tag Routine
29
Return

FIG. 3

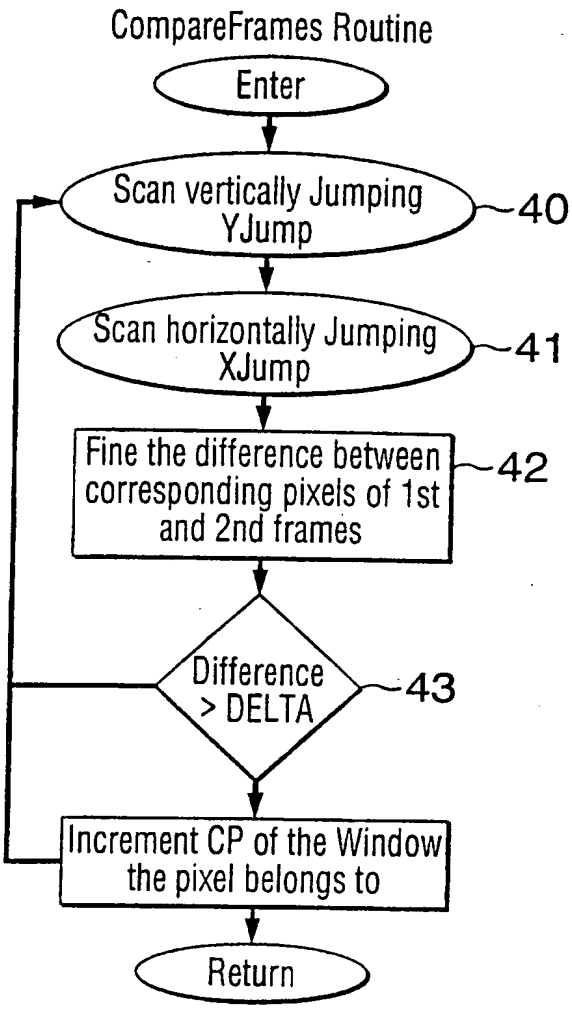

CompareFrames Routine
Enter
Scan vertically Jumping YJump
40
Scan horizontally Jumping XJump
41
Fine the difference between corresponding pixels of 1st and 2nd frames
42
Difference > DELTA
43
Increment CP of the Window the pixel belongs to
Return

FIG. 4

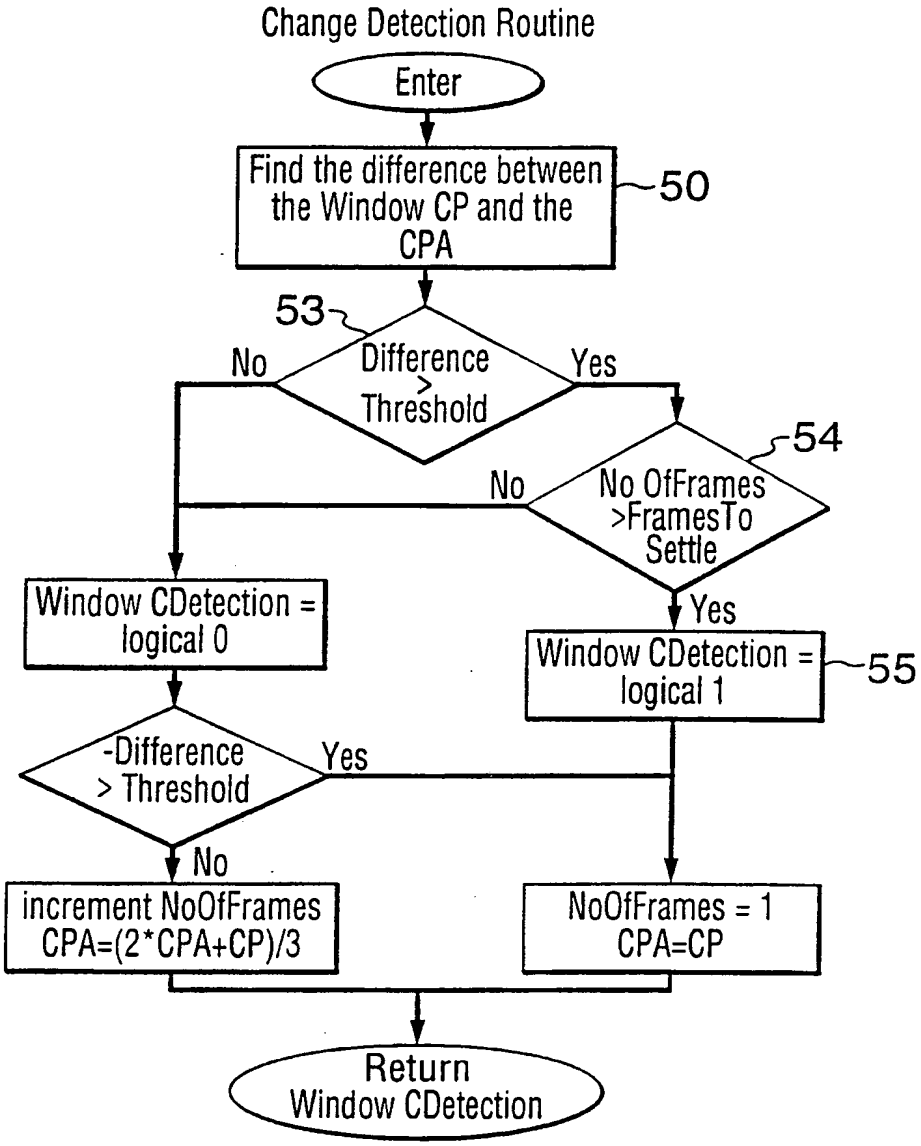

Change Detection Routine
Enter
Find the difference between the Window CP and the CPA
50
53
Difference > Threshold
No
Yes
54
No OfFrames >FramesTo Settle
No
Yes
Window CDetection = logical 0
Window CDetection = logical 1
55
-Difference > Threshold
Yes
No
increment NoOfFrames CPA=(2*CPA+CP)/3
NoOfFrames = 1 CPA=CP
Return Window CDetection

FIG. 5

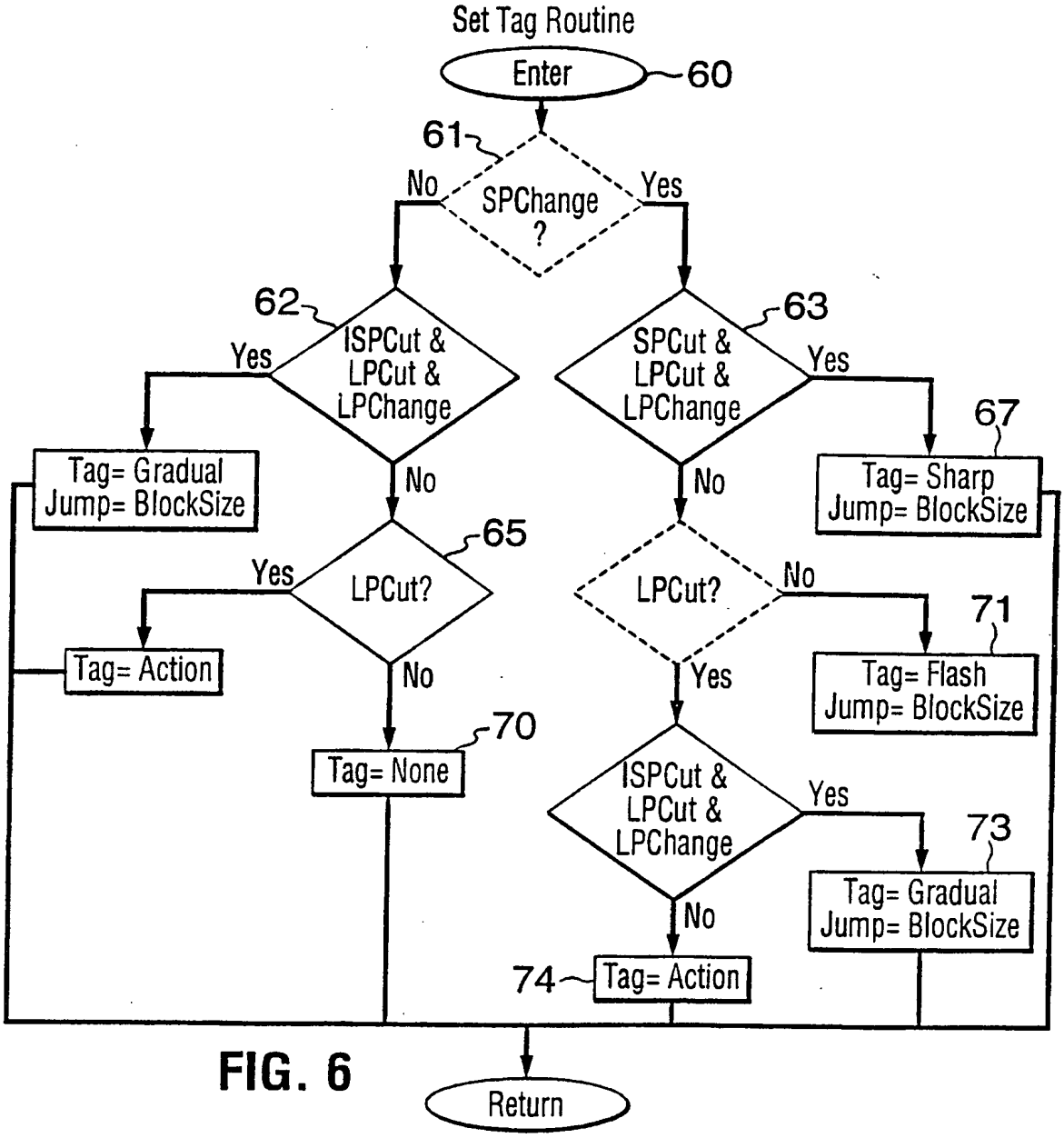

Set Tag Routine
Enter
60
61
SPChange ?
No
Yes
62
ISPCut & LPCut & LPChange
Yes
No
63
SPCut & LPCut & LPChange
Yes
No
67
Tag= Sharp Jump= BlockSize
Tag= Gradual Jump= BlockSize
65
LPCut?
Yes
No
Tag= Action
70
Tag= None
LPCut?
No
Yes
71
Tag= Flash Jump= BlockSize
ISPCut & LPCut & LPChange
Yes
No
73
Tag= Gradual Jump= BlockSize
74
Tag= Action
Return

FIG. 6

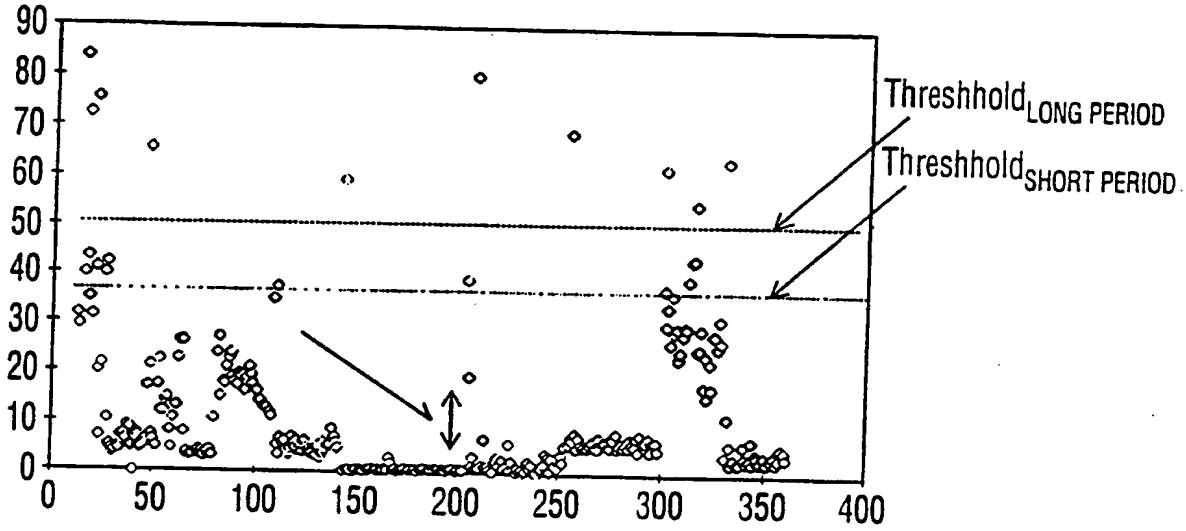

90
80
70
60
50
40
30
20
10
0
0
50
100
150
200
250
300
350
400
Threshhold LONG PERIOD
Threshhold SHORT PERIOD

FIG. 7

INTERNATIONAL SEARCH REPORT

Int. ational Application No
PCT/CA 97/00881

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04N5/14

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	EP 0 660 327 A (ATT) 28 June 1995 see page 7, line 20 - page 9, line 12	1,15 2-14,16, 17
X	EP 0 659 016 A (NTT) 21 June 1995 see page 3, line 27 - line 29	1,15
A	YEO ET AL: "rapid scene analysis on compressed video" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY., vol. 5, no. 6, December 1995, NEW YORK US, pages 533-544, XP000545960 see the whole document	1-17

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

2

Date of the actual completion of the international search	Date of mailing of the international search report
2 February 1998	12/02/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl. Fax: (+31-70) 340-3016	Yvonnet, J

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int...ational Application No
PCT/CA 97/00881

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 660327 A	28-06-95	CA 2133233 A	22-06-95
		JP 7203294 A	04-08-95
EP 659016 A	21-06-95	JP 8214210 A	20-08-96
		US 5642294 A	24-06-97